ABSTRACT:

In composite document image, handwritten and printed text is often found to be overlapped with printed lines. The problem becomes critical for obscure and broken lines at multiple positions. Consequently, line removal is unavoidable pre-processing stage in the development of robust object recognisers. Moreover, the restoration of the smash-up characters after removal of lines still persists to be a problem of interest. This paper presents a new approach to detect and remove unwanted printed line inherited in the text image at any position without character distortion to avoid restoration stage. The proposed technique is based on connected component analysis. Experiments are conducted using single line images that scanned and extracted manually from several documents and forms. It is demonstrated that our approach is equally suitable to deal with line removal in printed and handwritten text written in any language circumvent restoration stage. Promising results are reported in comparison with the other researchers in the state of the arts.